

OVERVIEW OF POPULATION VARIATION

Transition in: *After exploring the models of classical genetics, we are ready to return to populations to think about the frequency of traits within populations and the distribution of traits across populations. The phenomena and questions for this model return our attention to evolutionary thinking and to once again consider the role of natural selection in shaping the diversity we see on the planet today.*

NOTE: *Because this model occurs near the end of the year (in most sequences), teachers need to thoughtfully consider the number of days to allot to it, making sure to allow time for the speciation model and final return to Unity and Diversity at the end. Better to skip this model entirely than to never discuss speciation and descent with modification, and/or to run out of time before getting closure on the question of U and D. There are some suggestions in this overview for places where corners can be cut to save time, but as written this model should take about 5-7 days.*

Overview

We have explored the genetic and environmental factors that contribute to variation in individuals, but now ask students to step back and think about the distribution of variations within and across human populations. In developing a model to explain phenomena at the population level, we go deeper into the processes underlying natural selection and the distribution of traits in populations over time and space. Later, in the next model, we will extend this thinking to help us understand how new species develop. (This deeper understanding of natural selection ultimately positions us to address our overarching question from the beginning of the year about the origins of biodiversity and unity.)

We begin reorienting our thinking from the individual to the population level by looking at the occurrence of several traits within the class. (Note: if you did this activity at the beginning of classical genetics, you can revisit it and look at the data here.) Once we have a count of how many in the class have each phenotype, we ask students to think about what factors might determine which variation of a trait is more common. One idea that invariably surfaces is that more common traits are probably dominant. Our data allows us to easily test this idea. We show students which of the variants are dominant, and as often as not, the variation that is most common in the class turns out to be recessive. This allows us to quickly dispense with the notion that dominant traits are necessarily more common. Our class is a small population, but in classical genetics we saw an example from a large population - ABO blood types – so we revisit that here. We learned that the O allele is recessive, yet O is the most common blood type in the US. Other ideas students frequently mention as determining whether a trait is rare or common draw on our natural selection model – for example, students often propose that more common traits are more “advantageous”. If this idea comes up, discuss these particular traits in that light. Students will find that for many of the traits it is difficult to imagine how one variation might have an advantage over the other. We can speculate about some of the traits, but we can’t be sure, so this can remain as a question we need to investigate further. If this idea doesn’t come up here, that’s fine too – it will most certainly arise in the next learning segment. In any case, what’s important is that we recognize that we do not yet have all the ideas we need to explain why traits are more or less common in a population.

At this point we introduce the terms gene pool and frequency with a simple demonstration using squares of paper printed with either a 1 or a 2. Pick one of the traits we just observed, for example dimples or hand clasp. Ask students first, how many alleles they each have for the trait (hopefully they remember they have two!). Next, given their phenotype, what are their two alleles for this trait? Students showing

the recessive phenotype should remember they have 2,2. Give each of those students two squares with 2's on them. Students showing the dominant phenotype will (we hope!) know that they have a 1, but that the other allele could be a 1 or a 2, we can't know for sure. Give each of those students a square with a 1 on it. You can have an interesting discussion here about the chances of their second allele being a 2 vs. a 1 (for example, would most people be more likely to be 1,1 or 1,2? Why?) but ultimately you decide, for the purposes of the demonstration, how to allocate the second allele. So for example "let's say half of you (or 2/3 or whatever) are 1,2 and the others are 1,1." Pass out the second alleles accordingly. Now you can discuss the concepts "gene pool" and "frequency" with a concrete example, and take the conversation as far as you want. You can have students put their alleles in jars representing the male and female gene pools. You can talk about chances of different combinations in the next generation, previewing concepts that will come up in the simulation later in the sequence, such as "random mating with regards to a trait", "allele frequencies", and the connection between allele frequency and possible outcomes in the next generation.

Now that we have done some initial thinking about traits in populations, we look at maps showing the geographic distribution of phenotypes for several human traits – Huntington's disease, height, lactose intolerance, cystic fibrosis, and sickle cell anemia. We ask students first what they notice, and then what questions they have after looking at the maps. We hope they notice that in each case there are differences in the frequencies of phenotypes across different areas. We hope to eventually arrive at a driving question along the lines of, "What determines frequencies of phenotypes in a population, and why do the frequencies differ across different populations?" We elicit student ideas about the answer to this question more formally now, and generate a class list of initial model ideas.

To get a better understanding of how gene pools behave over time, we introduce an activity that uses different colors of beans to represent alleles in a gene pool. We explain that this simulation activity will serve a couple of purposes: it will make some of our new vocabulary more tangible, and it will be useful down the line in illustrating how populations behave and helping us answer our driving question. The procedure of the lab rests on two assumptions that we need to make very clear to students: 1) mating is random with regard to the trait, and 2) all individuals born survive and reproduce (in other words, there is no selection). The observations we make in the "bean lab" lead us to establish our first two model ideas about populations: *1) the frequencies of phenotypes in a population reflect the frequencies of alleles for the trait in the gene pool, and 2) if there is no selection, the frequency of alleles in the gene pool of a population does not change from one generation to the next.*

In debriefing the bean activity, students are asked to hypothesize what would happen to gene frequencies between generations if one of the combinations (say RR) produced a lethal condition. Students readily predict that the frequency of R would go down, and W would go up. At this point we look at the example of sickle cell anemia. Students do a reading on the disease to learn about its symptoms and its genetic basis (or use the information in the slides and discuss as a class). Given that sickle cell anemia is often fatal, we discuss what we would expect to happen to the frequency of the sickle cell allele in a population over many generations. Students should predict that, as we discussed in the bean lab, over time the frequency will go down, eventually approaching zero. Have groups think about how they would redesign the procedure of the bean lab to reflect this scenario, letting red beans represent HbS alleles, and white represent HbA. Their designs should reflect that many RR individuals die before reproducing, so they do not go back into the gene pool. If time allows have students try out their design. If time is short, have them predict what the data would show after two generations of selection against RR. Ideally they expect the frequency of R to get lower and lower, while W will become more and more common. Discuss, "Will

the R allele ever disappear completely?” Students usually recognize that even when the frequency gets very low, the allele is still present in heterozygotes so it will never be completely gone. As a result of our discussion, we add these ideas to our model: *Frequencies of alleles change over time when there is differential survival of phenotypes. The frequency of alleles associated with disadvantageous phenotypes decreases (but usually never disappears completely because of heterozygotes), and the frequency of advantageous alleles increases.*

Now we take a closer look at the map showing distribution of the sickle cell allele. What is surprising? Hopefully students notice that not only does the frequency vary dramatically geographically (even in areas that are close together), but that in some areas there is a surprisingly high frequency of the allele, given its lethal nature. Elicit student ideas about this. Direct their attention back to their initial ideas, particularly those that connect to natural selection. Hopefully they suggest that differences in the environments might explain this. At this point, show students the map of distribution of malaria and how it coincides with the distribution of the sickle cell allele. We explain a little about malaria, and the number of people it kills in the areas where it is prevalent. Discuss the connection between sickle cell and malaria resistance. You might show this excellent 15-minute video on how the connection between these two diseases was discovered. <https://www.biointeractive.org/classroom-resources/making-fittest-natural-selection-humans>.

So how might this connection with malaria affect the selection pressure on the sickle cell allele in high malaria areas? Depending on time available, there are several ways that students might grapple with this question. You can have them do the provided second extension to the bean lab, or, even better, lab groups can design an extension to the bean activity themselves that addresses this question. Their designs should somehow include selection on both RR and WW individuals: RR individuals die of sickle cell anemia and do not survive to reproduce, and many (but not all) WW individuals succumb to malaria, so they too die before reproducing. If time allows, students can actually try out their plan and see what happens. If time is shorter, you can simply discuss what they would predict, and possibly show data from a simulation, or from a previous year. Either way, this activity should lead students to conclude that in high malaria areas, there is selection against both RR and WW, so the heterozygotes (RW) have the greatest survival advantage (they are the “fittest”). As a result, the frequency of the R allele initially goes down, but instead of becoming vanishingly small as it would in areas with no malaria, it stabilizes at a moderate level.

After this activity it is time to return to the class’s initial model ideas. How do students feel about these ideas now? First we want to tease out ideas remaining that relate to natural selection – those that mention advantageous traits, and/or the role of the environment for example. Discuss how we now feel about these ideas, and how we might incorporate what we learned from the sickle cell anemia example into the model. Hopefully from this discussion we get model ideas that connect environmental factors to fitness, allowing us to explain why different variations might be more common in different areas. Student wording will vary, but something along the lines of: *Environmental factors determine which phenotypes are advantageous or disadvantageous, so distributions of phenotypes often differ geographically.* Although it isn’t necessary, students may also want to include a model statement about selection that favors heterozygotes, as in sickle cell: *In an environment where heterozygotes have an advantage over homozygotes, the frequency of a disadvantageous allele may remain unexpectedly high.*

We want to finalize the model at this point, so any initial ideas that we haven’t discussed previously should be addressed now. One idea that often comes up is mutation. Discuss the role of mutation in the sickle cell example. Students may remember from the reading that the sickle cell allele is the result of a

specific mutation in the gene that codes for hemoglobin. Mutation, in other words, is what initially created the new allele. Do we need this in our model? We already have the idea that new alleles are created by mutations in our inheritance model, but students may recognize that where a mutation initially occurred affects the distribution we see today, so they may want to include it in this model too.

Another idea that students often suggest is that migration (or simply “moving”) is a factor affecting the frequencies of phenotypes in different areas. Even if it is not suggested by students, this is an idea we want in the model, and we can accomplish that by looking at the example of sickle cell anemia. Looking at the map, where do students think the original mutation that created the allele might have occurred? If the mutation happened in Africa, why is it now on other continents, especially places like the U.S. where there is no malaria? The “migration” to North America, is of course an ugly chapter in our history and a sensitive subject, but one we need to acknowledge, and specifically to consider in explaining this phenomenon.

Once the model is finalized, we return to our driving question. Students write answers to the driving question using the specific example of sickle cell anemia. If time allows, students share their answers in groups and come up with a consensus “best” explanation.

***Transition out:** Now that we understand how frequencies of phenotypes and alleles across different population become different, we ask how such differences over time might lead to the rise of new species. The next model will explore the mechanisms of speciation, and inform our understanding of biodiversity. This will position us to finally return to the big question of Unity and Diversity.*

Overall Time: 5-7 days (depending on whether or not students design their own extension to the bean lab to model selection in sickle cell anemia, and whether they actually try out the procedures they design.)